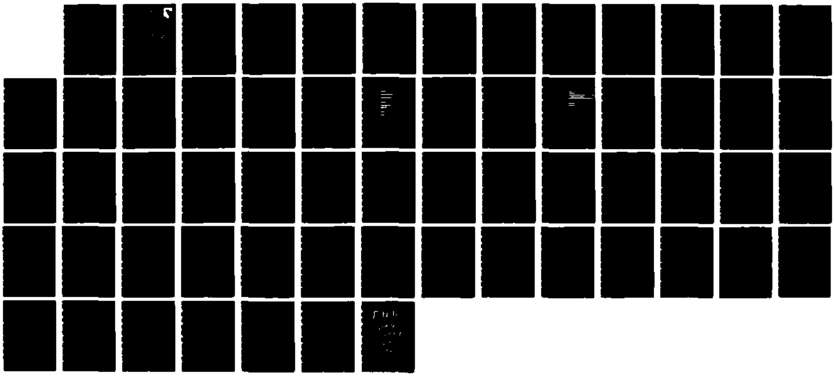
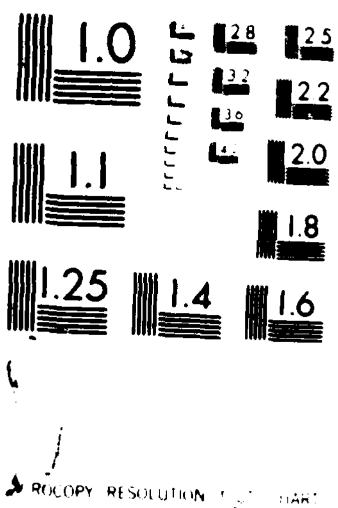
AD-A185 988 HUMAN PERFORMANE DATA NEEDED FOR TRAINING DEVICE 1/1
DESIGN DECISIONS(U) MACAULAY-BROWN INC FAIRBORN OH
G A KLEIN ET AL MAR 87 86-86.1-F AAMRL-TR-87-010
UNCLASSIFIED F33615-82-C-0513 F/G 5/9 NL




1.0
2.8
2.5
3.2
2.2
3.6
2.0
1.1
1.8
1.25
1.4
1.6

12

DTIC FILE COPY

AAMRL-TR-87-010

AD-A185 988

HUMAN PERFORMANCE DATA NEEDED FOR TRAINING DEVICE DESIGN DECISIONS (U)

GARY A. KLEIN
CHRISTOPHER P. BREZOVIC

KLEIN ASSOCIATES INC.
800 LIVERMORE STREET
P.O. BOX 264
YELLOW SPRINGS, OHIO 45387

DTIC
ELECTE
NOV 20 1987
S D
D

MARCH 1987

Approved for public release; distribution is unlimited.

HARRY G. ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY
HUMAN SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433-6573

87 11 3 489

NOTICES

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Please do not request copies of this report from Armstrong Aerospace Medical Research Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

Federal Government agencies and their contractors registered with Defense Technical Information Center should direct requests for copies of this report to:

Defense Technical Information Center
Cameron Station
Alexandria, Virginia 22314

TECHNICAL REVIEW AND APPROVAL

AAMRL-TR-87-010

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

CHARLES BATES, JR.
Director, Human Engineering Division
Armstrong Aerospace Medical Research Laboratory

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for Public Release; Distribution is Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) 86-86.1-F	5. MONITORING ORGANIZATION REPORT NUMBER(S) AAMRL-TR-87-010
6a. NAME OF PERFORMING ORGANIZATION Klein Associates Inc.* / 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION AAMRL/HEA
6c. ADDRESS (City, State, and ZIP Code) 800 Livermore Street, P.O. Box 264 Yellow Springs, OH 45387	7b. ADDRESS (City, State, and ZIP Code) Wright Patterson AFB, OH 45433-6573
8a. NAME OF FUNDING/SPONSORING ORGANIZATION / 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F33615-82-C-0513
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. 62202F; PROJECT NO. 7184; TASK NO 26; WORK UNIT ACCESSION NO. 06

11. TITLE (Include Security Classification) HUMAN PERFORMANCE DATA NEEDED FOR TRAINING DEVICE DESIGN DECISIONS (U)

12. PERSONAL AUTHOR(S) Klein, Gary A. and Brezovic, Christopher P.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
FINAL	FROM 1SEP85 TO 4APR86	1987 MARCH	56

16. SUPPLEMENTARY NOTATION
*Subcontract with (MacAulay Brown, Inc.) #1054-10B

17. COSATI CODES FIELD	GROUP	SUB-GROUP
05		
08		

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Training Device Design; Human Performance Literature; Design Decisions; Human Factors Data; Decision Strategies; Design Expertise

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The goal of this project was to study the types of human perception and performance information that training device designers need in making design decisions. There were two objectives: (1) to identify the types of human performance data needed to make these decisions and (2) to make recommendations about decision support for training device designers. A total of 50 experienced designers were studied. For a subset of 42 of these designers, the interviews focused on critical design decisions where there was a need for perception/information processing data.

Several types of data were collected from these interviews. One finding indicated that for the sample of critical decisions studied, systematic decision-making strategies were used in a minority of cases. A second result was the identification of frequent questions about human performance data. A third finding revealed the heavy reliance on informal experiments and analogous cases for guidance in resolving design questions, and the lesser reliance on published literature. Implications were presented for the

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Kenneth R. Boff	22b. TELEPHONE (Include Area Code) 513/255-7596; 22c. OFFICE SYMBOL AAMRL/HEA

DD Form 1473, JUN 86 Previous editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

Abstract continued:

development of a Designers' Associate. Future research areas were presented, primarily with methods for helping designers to extrapolate from research data and from analogous cases.

PREFACE

The authors wish to acknowledge Ken Boff for his support and encouragement and helpful criticisms throughout the effort. William Rouse and Edward Martin assisted greatly in arranging for interviews, Kathy Martin provided much help in conducting and managing the research, and valuable editorial comments were made by Janet Lincoln, Helen Klein, Janet Taynor, and Paula John.

Accession For
NTIS
By
Dist
A-1

TABLE OF CONTENTS

PREFACE..1

I. INTRODUCTION...3

II. METHODS..6

III. RESULTS AND DISCUSSION................................9

IV. CONCLUSIONS...30

REFERENCES..39

APPENDIX A Questions of Human Factors (HF) Literature......40

APPENDIX B Design Questions...............................50

APPENDIX C Recommendations................................55

I. INTRODUCTION

Just as equipment comes with a set of specifications to describe components, voltage limits, temperature tolerances, and acceptable interfaces with other pieces of equipment, so should there be a human use specification, with appropriate reference materials for design engineers. Because all equipment must interface at some point with a human operator, user, or maintainer, designers must understand the range of human tolerances. In some ways the story of Human Factors research is the preparation of this human specification. However, the task is not an easy one. Not only must research address the many unanswered questions pertaining to human performance, but the results must be communicated in a way which is meaningful and useful to designers.

The Integrated Perceptual Information for Designers (IPID) Program of the Armstrong Aerospace Medical Research Laboratory (AAMRL) is a comprehensive attempt to provide a human specification for design engineers.

The research effort described in this report was initiated as part of the IPID Program to study how training device designers make key decisions and to examine the type of human performance data they need for their decision making. The goal of the research was to ensure the compatibility of the IPID products and designers' needs.

After briefly reviewing the IPID Program, we will present a preliminary description of designer decision making and the results of the present study.

The IPID Program for Describing Human Characteristics

The three-phase IPID Program was initiated by the Harry G. Armstrong Aerospace Medical Research Laboratory to prepare this specification. The first stage of this effort was the preparation of the *Handbook of Perception and Human Performance*, a multi-volume work developed by eliciting summaries of the state-of-knowledge from leading researchers in the fields of human perception and performance. This material was written with the Human Factors and psychology specialist in mind.

The second stage of IPID is the preparation of the *Engineering Data Compendium*, another multi-volume work that is being written for the design engineer. The *Engineering Data Compendium* is being developed from selected material in the Handbook of Human Performance, judged to be relevant to the needs of designers. This is the IPID product intended as a human specification, a description of the various perceptual and performance characteristics of humans that affect our interfaces with hardware and software. It is carefully human factored and includes many tables of contents, indices, and cross references, allowing users to quickly find searched information. In addition, the entries are written to fall on two facing pages, carefully and clearly organized so that the most important relationships are highlighted. One of the many ways that the Engineering Data Compendium will stand as a model is in the strategy used to present information.

The third phase of IPID will be the conceptual development of a Designers' Associate. This is planned as an automated decision support system for design engineers, using the material in the Engineering Data Compendium as a partial data base, adding additional information, and overlaying this with intelligent support strategies for helping designers clarify their questions and focus their information search strategies.

The first of the three IPID phases is essentially completed. Phase I was established to provide information for the second and third phases of IPID. In addition, the Phase I focus was on training device designers, identified as a prime audience for the IPID products. The second phase is nearing completion (in terms of preparation for publication). Planning for the third phase, the Designers' Associate, is just beginning.

Goals of the Present Research

The major emphasis of this project is to learn how designers use human performance information to make decisions. We especially want AAMRL to be able to use the information collected in developing the Designers' Associate. This project has the greatest potential for synthesizing a strategy tailored to designers' needs.

The present research project is an attempt to answer two questions:

(a) What types of human performance data do training system designers need? We must both identify and understand their information needs in order to determine the contents of a decision support system. In addition, we wanted to learn how they currently obtain these data and evaluate their satisfaction with the methods they currently use to obtain these data.

(b) How can the IPID products be tailored to best meet the needs of designers? We were specifically interested in the Engineering Data Compendium and the Designers' Associate. We wanted to identify potential barriers that training device designers might encounter in trying to use these IPID products, as well as the ability to infer recommendations from the way designers make decisions and the types of data they need.

II. METHODS

Subjects

The subjects were 50 professional system designers employed at five private industrial firms and three separate Department of Defense (DoD) installations. Selection was based on two primary criteria: (1) employment as a training device design specialist and/or administrator and (2) several years' experience in that capacity. Our objective was to interview the most experienced designers available. Most often, subjects were scheduled for the interviews by the higher level management of the parent organization after preliminary discussion with someone representing or affiliated with AAMRL or with the Aeronautical Systems Division (ASD) of the Air Force Systems Command (AFSC).

Materials

A semistructured interview guide was used in the interview sessions. Biographical information and descriptive information, including the subjects' typical involvement in the design process at the organizational level and subsystem component level, were recorded in accordance with the guide outline.

Procedure

An interview session usually lasted about two hours and generally followed a three-phase outline. We prefaced the data collection phases with a short introduction to the *Engineering Data Compendium* and a description of our study, emphasizing that our interest was in the design decision process and the human performance data used in that process.

a) The first phase of the interview was the collection of biographical information and descriptive data on the subjects' design specialties and duties.

b) The second phase was to probe design decisions. We have recently developed a Critical Decision (CD) method as a knowledge elicitation strategy (Klein, Calderwood, & Clinton-Cirocco, 1985) and applied this strategy in the present study. We focused on the designers' recall of decision processes, asking designers to recount from memory a recent program effort. We probed the sources of information they used, the perceived utility of these sources, and the barriers to their use.

During their descriptions, we attempted to identify the most relevant of their decision points (those related to human performance questions and the need for more information). Specific questions were directed toward eliciting designer usage of human performance data.

c) The third phase of the interview consisted of two parts and involved a more detailed examination of the Engineering Data Compendium with the subjects. For the first part, we devised an exercise which involved searching the Engineering Data Compendium for data on a problem condition. Before each exercise, we explained to the subjects that the objective was to illustrate how the Engineering Data Compendium could be used by designers in the field. The search task commenced with the test question, "What are the effects of vibration on legibility?" The subjects were then invited to cross-index the Engineering Data Compendium as they would any research source.

The most relevant data were predetermined to be in entry 4.[illegible] of the Engineering Data Compendium: "Effect of Character Subtense on Display Legibility during Vibration." When subjects exhibited difficulty in locating this entry, we directed their search and continued to record any comments or suggestions the subjects made.

In the last part of the interview, future plans to configure the Engineering Data Compendium as an automated data base were described to the subjects. They were asked to identify features they would most like to see; they were also asked to identify presently available systems that any proposed design should emulate. This last phase of interview sessions was, on occasion, omitted or abridged when the earlier phases of the interview extended past the allotted two hours.

Two group interviews (involving nine of our subjects) and four individual interviews were informally conducted, producing no actual design project problems. However, these interviews provided important feedback on how human performance data are typically used, perceived gaps in the published literature, and on the Engineering Data Compendium. (This information was retained for part of the following analyses, where all 50 subjects are noted as comprising the data base.)

To summarize the procedure, a typical interview was conducted as follows: The senior author and a second interviewer met with the designer for a two-hour block of time. Demographic information was collected. The designer was asked to identify a recent design decision for which he or she needed to know something about human performance characteristics. We specifically looked at alternative answers and strategies the designer used to resolve the question. We were able to identify and probe several additional QDs with this subject. Next, we conducted a trial exercise in which the designer attempted to answer a preselected question using the Engineering Data Compendium. Finally, we asked about features to include and exclude in an automated data management system.

III. RESULTS AND DISCUSSION

In analyzing the interview data, two major topics were addressed. The first was a description of our subjects and their needs. What was the training and experience of our subjects? What did they need to know? How did they try to find it out? How useful were Human Factors data?

The second topic was an assessment of the IPID products. How easily did our subjects use the Engineering Data Compendium? How should the Engineering Data Compendium be distributed? What suggestions did our subjects make for the Designers' Associate?

Characteristics and Human Performance Information Needs of Training Device Designers

A. Who were the subjects: Demographics

The 50 subjects interviewed were experienced training device designers. Nearly 60% of the subjects had been awarded engineering degrees, most in electrical and/or mechanical specialties (See Table 1). The non-engineering degree subjects were diverse in their academic backgrounds, with the two most frequent fields Behavioral/Social Science and Computer Science. Table 1 shows the average years experience of the subjects from each discipline. The overall average was 9.5 years of experience. It should be noted that Table 1 represents data for only 45 subjects, primarily those where we probed Critical Decisions. The remaining five subjects in our sample were interviewed either as part of a group or informally; little or no demographic data were collected from these subjects.

The academic backgrounds of our subjects were as follows: 11% Ph. D. degrees, 42% MA/MS, 47% BA/BS. Most of them worked in industry (72%) rather than in the DoD (28%). Most were not Human Factors specialists: 22% vs. 78%.

The subjects were asked to note their primary and secondary training system specialties. The most frequent primary and related areas of training system design were in displays, hardware, and instructor/operator stations (IOS), in decreasing level of frequency. For their typical level of effort at specific stages in the design process, the subjects were asked to respond with 'high,' 'medium,' 'low,' or 'none' when shown the preselected stage categories. The subjects were encouraged to mark all

Table 1

SUBJECT DEMOGRAPHICS

Discipline	Frequency	Avg. Yrs. Exp.
Engineering		
Electrical	11	14
Mechanical	5	6
Elec & Mech	2	7
Aeronautical	6	13
Industrial	4	8
Design	2	21
Systems	1	--
Reliability	1	.3
Non-Engineering		
Behav/Soc Sci	4	9
Comp Sci	4	10
Math	3	5
Mgmt	2	17
Bio/Phys/Med	1	4

entries that applied. Tables 2 and 3 illustrate the frequency distribution of the responses.

(b) What did the subjects need to know: Design Questions

Perhaps the most important data in this study are the design questions for which our sample of training device designers needed more information. Any Engineering Data Compendium or Designers' Associate must try to provide these types of data. We identified 135 unique questions. These are listed in Appendix A.

In order to bring these data into better focus, we classified the 135 design questions into more general categories that were suggested by the data. Appendix B presents a listing of specific design questions organized by category. In Table 4 we have presented our classification along with a frequency distribution.

Table 4 shows that most of the questions were about visual perception, primarily related to CRT usage. Designers had to decide how to select letter and character sizes and portray symbols on CRTs; these

Table 2

FREQUENCY OF DESIGNERS' REPORT OF SYSTEM SPECIALTY AREAS

Training System Specialty	Relevance to Job Requirements	
	Primary	Related
Displays	20	9
Hardware	16	9
Software	9	9
Computer Interface	6	10
Instructor/ Operator Station (IOS)	10	12
Motion Systems	1	0
Aircraft Modeling & Flight Control	1	0
Computer Aided Instruction	1	0
Other	9	2
Total	73	51

Table 3

FREQUENCY COUNT OF DESIGNER REPORTS OF PERSONAL INVOLVEMENT IN THE DESIGN PROCESS

Phase of Design Process	Level of Involvement	
	High	Medium
Identifying/ Specifying a Training need	15	9
Conceptual Design	23	7
Preliminary Design/ Functional Spec	21	8
Detail Design	16	9
System Test & Evaluation	8	12
Sustain Engineering Change Proposals (ECPs)	7	9
Totals	90	54

about anomalous CRT effects, e.g., eye strain, Moire effects, etc. In addition to CRT issues, designers want to know more about night vision and the effect of Training Device (TD) layout on visual perception. Next to vision, the most frequently queried domain in this sample was motion perception. Some of the major issues were presentation and detection of motion cues and tolerance for motion forces. The last major topic was the area of controls--sensitivity, spacing and layout, different options, and safety.

(c) What data sources do TD designers use?

Having identified the most frequent content questions, we turned to an examination of the types of data sources our subjects attempted to use in answering these questions.

The major focus of our efforts on this issue was the subset of design questions that were encountered on actual projects. Only 40 of our subjects produced specific design cases, and of these only 37 responded to our probes specifically enough to allow data analyses. We were particularly interested in the problem-solving strategies used by designers. What sources did they use and how satisfactory were they? We recorded 76 decision points and 181 sources in response to the probes. From these, we identified three major classes of data sources and several subclasses.

The classes were (1) professional background or actions suggested by the designer's own experiential judgment, (2) query of informed personnel, and (3) technical literature.

We reviewed the decision accounts and the designers' comments on the above resources to appraise the utility of the different types of information. Five rating categories were used. For each incident, a data type may have been key to solving a problem, or it may have helped to solve it, or it may have been used to some small extent, or it may have been useless. Finally, the data type may have been searched for without success. Figure 1 shows the frequency of instances where a source of information either directly answered a question or helped to answer the question (the first two of the five categories of utility).

Table 4 - Categories of Design Questions	
Question Category	Number of Inquiries
Vision	
CRT	
general	14
alphanumeric	5
symbols	18
dynamics	10
organization	6
anomalous effects	14
color	5
Night Vision	6
Vision and TD Layout	12
Visual vs. Vestibular Cues	1
Motion	
Presentation	6
Detection	6
Audition	2
Workload	12
Controls	
Sensitivity	4
Spacing	1
Input Methods	7
Safety	1
General	4
Anthropometry	1

(1) The professional background category included 3 subcategories: experiential, analogues and prototypes, and mock-ups and quasi-experimentation.

The experiential judgment category is obviously involved every time a designer works on a system. It includes two types of experiential bases, perceptual experience by the designer who may have used the same or similar field equipment to be simulated and experiential judgment with the TD technologies.

Some examples will help to illustrate this subcategory. In one, the designer had to specify the visual cues used by helicopter pilots in night/day attack missions. Since the designer was a veteran pilot of helicopter warfare, he imagined himself piloting a helicopter in particular situations and was able to adjust the parameters of field-of-view, luminance, etc. In another case, a designer found needed results on g-force tolerances produced by standard laboratory study, but was suspicious of their low range of values. He investigated further and found that naive subjects were used in the experiments. Experienced pilots could handle much greater accelerations. He relied on his experience rather than on the data. Subsequent research with experienced pilots vindicated his judgment.

Analogue and prototype resources were those devices or systems specifically identified by subjects as similar or familiar projects they had previously encountered. By identifying these similar cases the designers were able to apply the results and methods to their own questions, making modifications where necessary. This subcategory accounted for a large percentage of solutions--20.3%. It appears to be a standard and effective means of handling design questions. For example, a designer remembered a prior maintenance TD that was fielded and had caused confusion for trainees who expected all hydraulic lines to be color coded as in the TD. He, therefore, decided to include pictures of the actual equipment in the new TD. In another example, a designer faced the problem of specifying motion cue onset lags for a helicopter TD, which he solved by adapting values he had obtained on a previous helicopter TD project.

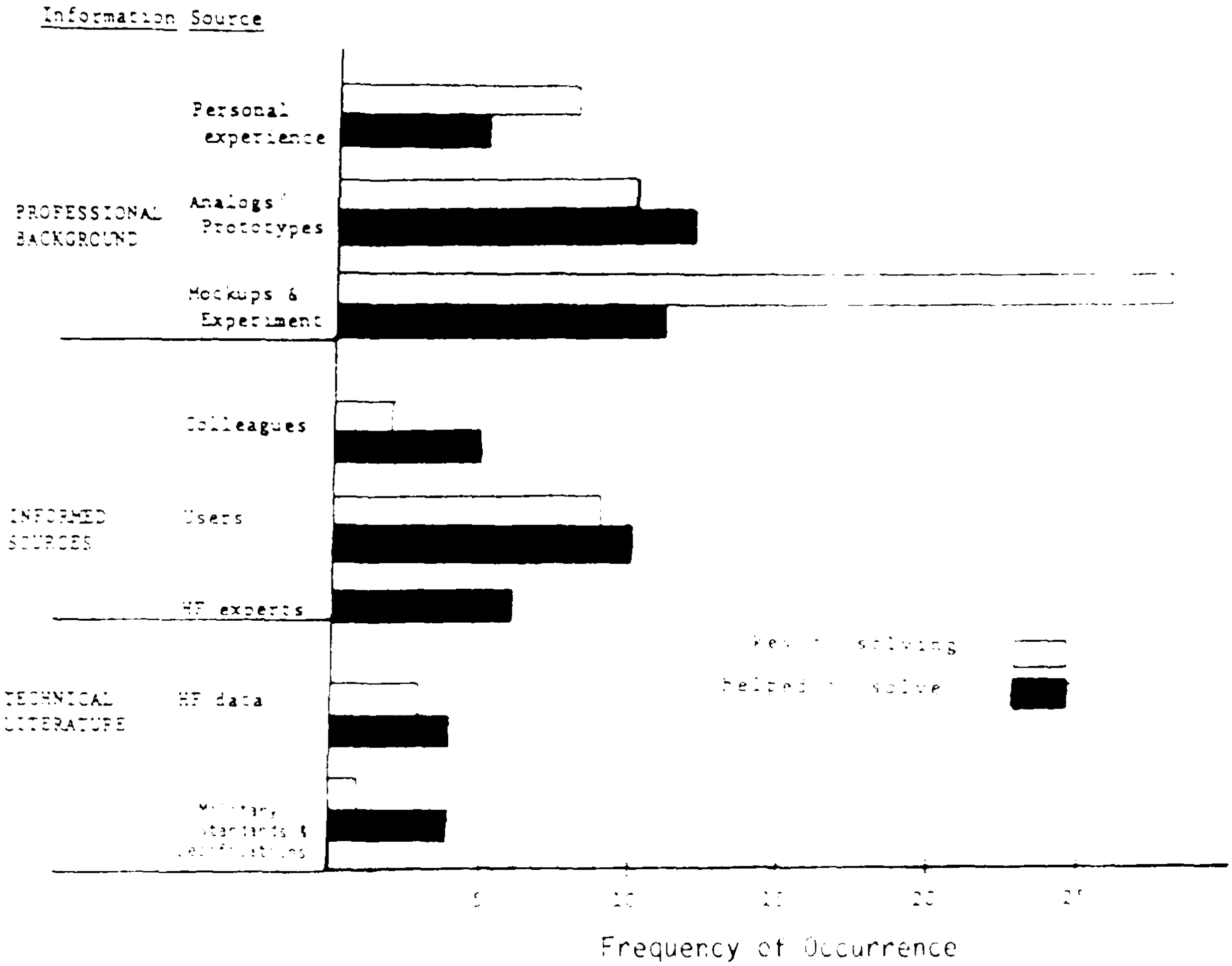


Figure 1: Value of Different Sources of Information

Informal experiments with mock-ups (including breadboard, brassboard, and prototype models) were the most common means of resolving questions. The quasi-experimentation efforts were routinely nonrigorous. The designers were interested in a general introduction to the phenomenon they were dealing with, in order to achieve 'ballpark' figures as opposed to accuracy and precision. This included sheer trial-and-error in some cases. Preference for this approach must be considered in light of one of the most frequent complaints expressed by the designers we interviewed: the attempt to engage staff HF experts in resolving specific design problems was frustrating. HF personnel would run controlled experiments that were both too time consuming and too limited in focus.

Some examples of the use of informal experiments may be helpful. Input and control mode configurations on a new IOS presented a complex problem to one designer who began the solution process by brainstorming with colleagues and then constructing a rough mock-up which he tested with the end users. In another problem, the question was how much resistance to build into one critical function of a dual mode switch. As the designer could find no relevant published research, he fashioned a solution by constructing a mock-up and conducting mini-experiments with a few end users. Another example was the mode of deriving a format for an auditory warning message to be placed in a jet fighter cockpit. Through experimentation with repeating/nonrepeating messages at different cue frequencies and decibels, a user-acceptable product was determined.

(2) The second general class of resources used by the designers was informed personnel. That is, the designers would query those identified as having some possible input into the problem resolution process.

Three sources were most mentioned: professional colleagues of the designer; users of the actual or similar field equipment; and Human Factors experts. Generally, all three subcategories were of fairly low frequency in our survey, especially professional colleagues and Human Factors experts. We were surprised by the scarcity of incidents where colleagues played a role, but Allen (1977) reported the same finding.

(3) The third general class of resources used by the designers was technical literature. This included two subcategories: HF articles (in

professional journals and technical reports from the government or private industry), and Military Standards (MIL STDs) and Specifications.

Figure 1 reveals striking contrasts in ratings of resource utility. The professional background class and the technical literature class had occurred with roughly similar frequencies but had markedly different assessed utility ratings. The professional background resources accounted for 68.5% of the highest ratings compared to 11.1% for the technical literature category. Mock-ups and experimentation accounted for 52.7% of the total cases where any particular piece of information was key to solving a problem. Human Factors data and MIL STDs or Specifications were about equal in providing solutions to the designers; 6.5% and 4.6%, respectively.

The general pattern of source utility supports the findings of Allen (1977) on the identification of information usage patterns of project engineers. The focus in this earlier effort was on the frequency of using information sources summed over the life of a project. He studied 17 technological projects in which the subjects were predominately engineers working on government projects. Although the methodology and orientation were different, Allen reported a pattern of results similar to ours. The sources that were identified as highly used in Allen's study are the same sources rated as helpful in our study. Matching his data to those reported in Figure 2, we find that personal experience accounted for 8% in Allen's study vs. 7% in our study; analysis and experimentation accounted for 31% in Allen's research vs. 22% for our design engineers. User inputs accounted for 19% in Allen's findings vs. 14.4% in our research; technical staff (HF experts in our study) accounted for 6% in Allen's work vs. 6.1% in ours. The only discrepancy involved technical literature: 8% for Allen vs. 26.5% in our research. Our findings may have shown a higher use of technical literature because of the nature of the critical decisions we probed (areas where human performance data were needed).

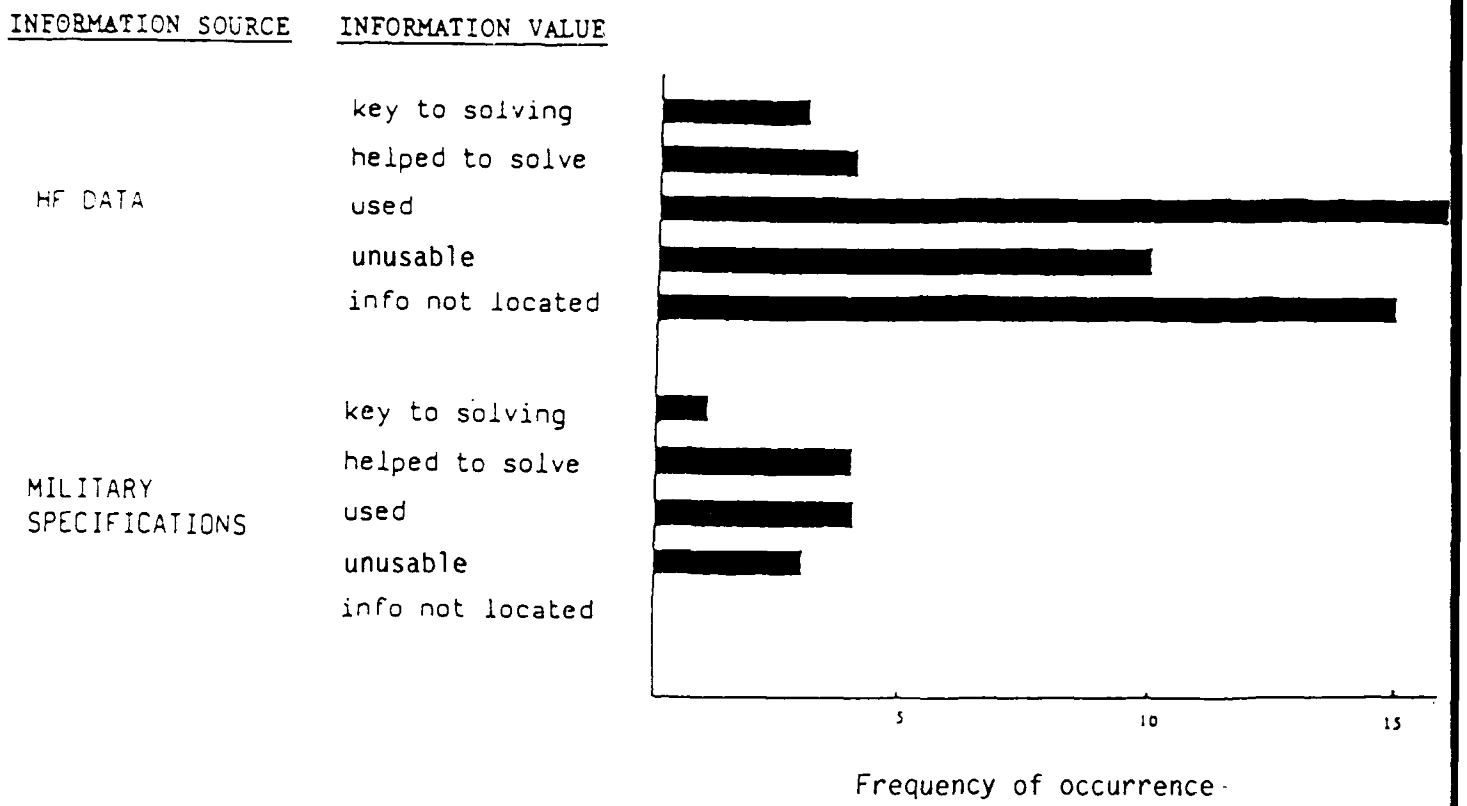


FIGURE 2: EVALUATION OF USEFULNESS OF TECHNICAL LITERATURE

<u>Information Value Scale</u>

key to solving: necessary information.

helped to solve: neither necessary nor sufficient.

used: information sought and applied but not essential.

unusable:: unable to apply information.

info not located: information sought but not found.

Because technical literature forms the basis for IPID, we studied resource in greater detail. Follow-up questions produced the results in Figure 2, which includes the full five category rating scale for utility.

The poor success by designers in attempts to apply Human Factors data is apparent upon inspection of Figure 2. Human Factors data were rated ineffective 20.8% of the time when this class was explicitly searched to help resolve design questions. More interesting is the fact that close to one-third of the time, HF data were judged simply not to be available. The other technical literature forms, MIL STDs and Specifications, were less frequently searched.

Our next question was to determine why technical HF literature was of such little use. For the sample of design questions we probed, we identified two main reasons for failure to use Human Factors data: difficulty in finding articles and difficulty in extrapolating results. After data collection was completed, we tabulated the number of times that each type of evaluation was made when Human Factors data were mentioned. The frequency data are presented in Table 5.

Table 5

PROPORTION OF DIFFERENT REASONS FOR USER DISSATISFACTIONS WITH HUMAN FACTORS DATA*

	Proportion
Difficult to find articles	(25.5%)
Lack of confidence in research methods and results	(0.0%)
Difficult to extrapolate from results:	
Information too general	(25.5%)
Irrelevant methods	(38.2%)
Conflicting data	(9.1%)

*Based on 76 probed critical design decisions

Access was a problem. In more than 25% of the cases, the subjects reported they could not locate relevant studies despite their searches. We also found a number of cases not reported in Table 5 where subjects did not attempt searches because the perceived probability of finding relevant data was too low or because the subject did not know how to conduct an effective search. An additional access problem was the difficulty in finding relevant information in the article.

Several points emerge from Table 5. 1) There were no incidents of any subject, whether HF specialist or design engineer, expressing any concern about the statistical reliability of Human Factors data. This stands in contrast to researchers involved in generating and evaluating reports of experiments, who continually worry about the reliability of group differences. Our users showed no concerns on this score, even for technical reports not exposed to the same level of scientific scrutiny as journal articles.

2) Another finding is that nearly two-thirds of the problems encountered in attempting to use Human Factors data occurred when designers tried to extrapolate from the information they found to specific projects on which they were working.

3) Over half of the extrapolation difficulties came from inappropriate or irrelevant methodologies. This problem typically arose when the experimental conditions differed from the operational design conditions. Designers did not know how to take these differences into account, finding only limited use for such data or rejecting the findings altogether as irrelevant. Two examples of this are: closure rate data obtained from studies on surface ships which could not be extrapolated to TDs for aircraft formation and refueling; and g-force tolerance limits gathered on inexperienced (college) subjects which were considered too low to be generalized to experienced fighter pilots. Designers found studies on specific stimulus effects in the literature, but their design problems concerned various combinations of effects that were not addressed.

(d) How did the designers make decisions?

Table 6 shows our categorization for the 76 probed critical decisions. There are two major categories in Table 6 and, in general, the categories are arranged in order from recognitional matches at one extreme of decision strategies to careful deliberations between options at the other extreme.

The first subcategory in Table 6 is for decision points wherein the designer recognized a match to a previous analogous case, or to a general class of cases. (We have called this latter situation a match to a prototype, using prototype in the cognitive psychology sense as referring to the synthesis of several instances.) This category showed the highest frequency, accounting for 22% of the cases. An example was a designer who had to specify the luminance for a CRT and realized that it would be the same as for a previous TD so he simply used that value. He did not attempt to identify other options, or to evaluate advantages and disadvantages of options.

The second subcategory involved serial matches. This was an extended form of recognitional matching; the initial match was not successful and

Table 6

CATEGORIES OF DECISION MAKING STRATEGIES SHOWN BY THE DESIGN ENGINEERS

	Frequency
Pattern Matching	
Match to Analogues/Prototypes	17
Serial Matches	14
Deliberation	
Mock-ups & Experiments	14
Decision Analyses	12
Adjustments	7
Other	8
Not Solved	4
Total	76

led to subsequent attempts to match. It is worth noting that there was still no reported attempt to identify and evaluate options by comparing one to another. If the first option had worked, there would have been no further effort. This subcategory accounted for 18% of the cases. An example would be a design team wrestling with the problem of how to represent and organize information on a CRT screen. They rely on mock-ups to test one potential solution after another until they find one that works.

The remaining subcategories involved careful deliberation: more than one option was considered at a time. One way this happened was through the use of informal and formal experiments and mock-ups to contrast alternatives. This occurred in 18% of the cases. The use of mock-ups here was different in that there was an intentional goal of contrasting alternatives. (For serial matching, mock-ups were used for the search to find an option that worked.) An example is a study examining different types of night vision goggles to see their impacts on TD operations.

The next subcategory was the use of decision analysis. This was a specific attempt to simultaneously contrast several options by determining strengths and weaknesses of each or by rating each along a set of pre-defined dimensions. This is the form of decision making that forms the basis of many prescriptive methods. We found this in almost 16% of the cases. Often it overlapped other categories (e.g., some options were generated through analogy, or the analysis moved to a certain point and then a hands-on study was performed).

We have included a subcategory of decisions labelled "adjustments." This occurred for approximately 10% of the cases. For example: how loud should a tone be in order to serve as an effective warning? These were carefully deliberated decisions, but they resembled continuous tracking tasks more than decision tasks in which subjects select between discrete options. Often the starting point was a match to an analogue or prototype, followed by trial-and-error using mock-ups.

Another 8% of the cases were deliberated but did not fall in these three subcategories, and 5% of the cases had not reached resolution at the time of the interview and were not categorized.

These findings support Rouse's (1983) observations about the differences between researchers and designers.

> While the objective-driven optimizer is likely to actively canvass a range of alternatives by searching the available knowledge base, the constraints-driven satisficer is unlikely to see any merit in such a strategy. Instead, new information will only be sought if the information at hand is unsatisfactory or, in some cases, if the new information is in a readily usable form and very easily accessed. The archival journal literature that comprises most of the research base by no means satisfies these conditions; even when available by online retrieval systems, this literature is oriented towards general understanding rather than solving specific design problems. Therefore, it is fairly clear why the practitioner is unlikely to access the research base in the course of pursuing design (Rouse, W. B., 1983).

Another observation is that although the people we interviewed had many years of experience, they tended not to have the level of expertise that would qualify them as true experts in the sense that chess grandmasters are. High levels of mastery allow people to immediately judge whether a new case is typical or atypical. In our research (Klein et al., 1985) we found that experts rarely used analogues to guide their behavior because their previous experiences had blended together to form prototypes. Our subjects were senior fire ground commanders who had experienced hundreds of examples of each type of fire. The designers had not encountered such a wide range of directly relevant cases. Rarely had any designer been experienced with 10 or more pieces of the same equipment. This would explain why the designers we studied relied more heavily on analogues than on prototypes. For the 76 decision points there were 19 direct mentions of analogues and 13 mentions of prototypes. (These include direct matches plus other related uses of analogues and prototypes.) The ratio of analogue to prototype use is much higher here than for fire ground commanders.

What description does this generate for experienced designers? In the standard case, the designer reaching a decision point would attempt to find a match to an analogue or prototype, and would likely try some informal research to obtain a better value, to fine tune the initial match. In some cases, there would be more than one option, because of recommendations from authority or from genuine uncertainty, and these cases would be handled in a simple fashion (e.g., the option with the lowest cost) or through a task analysis or experiment.

To the extent that an analysis is done, it follows along Tversky's (1977) concept of Elimination By Aspects. First, all options that are not feasible are eliminated, then options that are too expensive, and finally the cheapest acceptable option is selected. However, even this is too formal. Designers rarely looked at all options, and they attempted to "fix" options that were not working. They were developing options, not just generating and evaluating them.

Designers appear to avoid formal decision making, just as Janis and Mann (1977) said. Formal analysis is time consuming and effortful. There is no end to the analyses that can be performed. Designers simplify their task by considering only the most obvious option(s). They are less interested in adding to their option set than in working with the best option of the ones they have identified to make it as effective as possible. They play with the problem resolution to find a way of viewing the task that will make things obvious. A few examples will illustrate this point. A designer felt that it would be safest to use larger characters on a CRT and was not interested in evidence to the contrary. Another designer needed to portray oscilloscope wave forms to trainees and never really considered using CRTs; the designer had seen situations where training devices were rejected simply because of poor fidelity to field equipment and had no intention of repeating that type of mistake. In a final example, a government worker needed a specification for an Instructor/Operator Station, so he used and modified the last specification that went out in an RFP. There was no attempt to look further, to construct the ideal specification.

Is this a poor strategy? It is possible that a Designers' Associate could help designers to do better than to satisfice. However, many decision support systems have been developed based on formal methods of decision analysis, and have been rejected in the field. It is perhaps wiser to develop a decision support system that is compatible with the natural decision style of designers.

The training device designers in our sample were not trying to make revolutionary changes. Instead, they were seeking evolutionary changes, ways of applying previously developed technology in new areas. The last version of an approach was basically a test bed which illustrated potential Human Factors problems. Lacking an analogue to serve as a test bed, designers would turn to informal studies to create analogues. Designers never start with a tabula rasa.

Assessment of the IPID Products

The next area of investigation was the examination of the Engineering Data Compendium and Designers' Associate as useful products for TD designers.

(a) How easily did designers use the Engineering Data Compendium?

One of our main goals was to assess the usability of the Engineering Data Compendium. There were obvious limitations for such an assessment. At the time we conducted our research, only two of the subsections had been completed, so the vast majority of the materials could not be examined, let alone evaluated. Even if the entire Engineering Data Compendium had been completed, a thorough evaluation would have been a monumental task. Nevertheless, we attempted a limited evaluation as a means of generating hypotheses about the ways users would interact with this type of material.

We studied the applicability of the Engineering Data Compendium to design problems in two ways. 1) We conducted a trial exercise using a predetermined design question and targeting the relevant data, and 2) we examined probed CD points to determine where the Engineering Data Compendium would have been helpful.

The trial exercise provided feedback on format, perceived utility of the information, and confidence in future utilization of the Engineering Data Compendium. Consistently favorable comments were made on the physical layout of the information presented and the elimination of theoretical introductory material to the entries.

The critical remarks were generally of two types: criticisms of the incompleteness of the data for design purposes and suggestions for improving the organization of the document or for keeping the document updated.

The most frequent complaints were that insufficient information was included about the nature of the error measures and the characteristics of the subjects used. In our trial section on vibration and symbol legibility (Entry 4.112), some typical remarks were: "What does 'error' mean?; What constitutes performance criteria?; How important are errors for the training task?; Do the vibration ranges convert to g-forces?; How experienced were the subjects with sustaining performance under vibration forces?; How do the data apply to text strings?"

Other concerns the designers had were on the lack of information on confidence limits for the data and the comparison of these data to established MIL STDs. One designer noted, "These data are on thresholds--we need to have values not on 95% or even 99% accuracy but on 100% accuracy because that's the performance level we've got to have in the field. We have to design out the most amount of error we can."

Some questions showed how hard it was for our subjects to extrapolate from a study where subject type and experience are different from the training environment. This was described earlier in a case in which a designer noted that the g-force tolerance levels established in one laboratory study were for naive subjects and he was designing for veteran pilots. He recognized that these levels were too low and wrote to the experimenters advising them to rerun the study with experienced pilots.

A strong doubt voiced by another designer was the practical value for designers of the Engineering Data Compendium; specifically, that the interplay of natural context and any phenomenon of interest is so complex that research from laboratory settings could not be reliably generalized, or that atomistic findings could not be simply reconstructed additively to describe any larger complex phenomenon. The example he chose was the interdependence of luminance, contrast, and noise effects on cockpit displays when the external environment levels of each were changing.

On general layout, a common remark was that special guidelines would be necessary to help users locate topics they were searching for, as well as related issues that would be of interest.

Several Human Factors professionals remarked that technical terminology should have been avoided: "Why 'angular subtense'? Why not just 'visual image size'? Other engineers might need a Human Factors specialist to help interpret the data." And, in another designer's words, "Visual science people have got their terminology, (other) engineers have got their own jargon. It's difficult sometimes to go back and forth (between the different fields)."

The importance of keeping the Engineering Data Compendium current was repeatedly mentioned. It was suggested technology update sections be sent to users, particularly on computer interface issues such as different input modes (e.g., mouse, touchscreen, light pen, etc.) for different tasks.

The second method of studying the Engineering Data Compendium's utility in the field was to analyze how it would have been used on the critical decision points we identified on the design project examples collected. We found 31 out of 76 cases where the Engineering Data Compendium was clearly relevant. Such cases included identifying differences between monocular focal distance and focal distance during convergence, specifying character height on CRTs on "busy" backgrounds, and identifying the relationship between CRT refresh rates and image stability.

The remainder of the design decisions were not so clearly amenable to solution by appeal to any published human performance data. These design decisions for which the Engineering Data Compendium was judged not applicable primarily involved technology/field-specific issues. These included questions that arose over the impact of specific instrument or machine configurations on human perceptual ability or processing capacity. For example, what are the perceptual tolerances of different refresh rates of different CRTs in proximity? What are the effective vibration threshold cues to pilots of particular aircraft about to stall? Such system-based issues rarely are treated in available technical literature as system-specific findings. Rather, those entries that do exist are typically treatments of general considerations of the issue. The designers are therefore left to their own methods to resolve the research issues. Most frequently, these designers attempted to extrapolate from similar systems or to conduct experiments or interviews with end users. The Engineering Data Compendium does not contain the system specificity for its data that designers are looking for in such situations.

In addition there were some design questions for which the Engineering Data Compendium was judged simply not appropriate: a mixture of issues not addressed by the Engineering Data Compendium as presently structured (e.g., anthropometry and physiology) and unresolved issues (example: does the eye consistently exhibit a fixed saccadic suppression interval?).

(b) How should the Engineering Data Compendium be distributed?

The general preference for how to distribute the Engineering Data Compendium differed among designer specialties. For design engineers there was a preference for a central location of at least one set (e.g. the technical library or the engineering manager's office), plus additional sets to be located within access to a working group or branch. In some cases, it was judged sufficient to have one set per building. Only 11 of 28 subjects wanted a personal copy.

Human Factors professionals wanted more copies than other designers in the study. The average preference for an organizational distribution was one set for every 15 HF professionals, although the range was one set per every three to one set per 50 people. It was also more likely for the HF professionals to desire a personal copy; 9 out of 11 chose so.

Allen (1977) has studied the acquisition pattern of books and journals by engineers when at work on specific projects. This investigation revealed that engineers were most likely to access formal literature when the information was close at hand, more specifically, in their files or those of familiar colleagues. In fact, 92% of all formal literature acquisition occurred in this fashion for his sample. Furthermore, Frohman (1968, cited in Allen) has found that the extent of library use was an inverse function of the distance between a work group and the library. The implication for the distribution of the Engineering Data Compendium is that the volumes should be located as close to the working group as possible to improve usage of the document. If a set cannot be placed in the hands of the chief project engineer or the engineering manager, then suboptimal usage should be expected in direct relation to the distance of a set to the work group.

(c) Suggestions for the Designers' Associate

We used different questions to elicit suggestions about how to structure a Designers' Associate. Most of our subjects liked the idea of an automated data base. They felt that a chance to work with the system, as well as retrieve data from it, would be helpful. They liked the idea of a keyword querying capability and artificial intelligence for misspellings and synonyms. They also hoped to get caveats for applying data and recommended a case study addendum that would be a repository of past projects, including how things were done and what problems were encountered.

Beyond this, there were individual recommendations for window displays, a naturalistic query language, ways of varying guidance levels for novices and experts, and designated function keys. They advised the developers to avoid light pens and touch-sensitive and low-resolution displays. They appreciated the ease of updating an automated system but were concerned about the loss of resolution for a CRT and how this would degrade the graphics of the Engineering Data Compendium. They recommended the use of trees for easier search patterns. All of the suggestions were tabulated and reproduced in Appendix C. However, our subjects were not experienced in designing expert systems, data banks, reference materials, or any of the other aspects of the Designers' Associate. In general their suggestions did not include any grand schemes or novel ideas.

IV. CONCLUSIONS

In carrying out this project we learned many things about training systems design and, more specifically, about design decisions. In this section we will (a) trace some implications for the Engineering Data Compendium, and (b) do the same for the Designers' Associate. We will then (c) examine topics for future research.

(a) Implications for the Engineering Data Compendium.

Potential users overwhelmingly responded positively to the prototype shown to them. The thought and professionalism that went into the development, organization, and layout of the Engineering Data Compendium were recognized. This study did not uncover significant problems with the Engineering Data Compendium.

We conducted the trial exercise to test the use of the Engineering Data Compendium and found that virtually every subject was able to rapidly find the most applicable entry. The organization of the Engineering Data Compendium is quite clear and easily understood. Some subjects raised issues about jargon and metrics, but these will always be problems since there is no one standardized approach. Different fields use different terms, and it is not always useful to rely on neutral language since this may confuse professionals who have become used to jargon.

Table 4 (see page 14) can be viewed as the basis of a supplementary guide for training device designers, helping them to find the topic of greatest relevance (e.g., CRT alphanumerics, night vision, motion cues). Such a supplement might include all the entries relating to each of these topics, thereby simplifying the search strategy.

A problem that may be inherent is the relationship between the Engineering Data Compendium presentation and designer problems. For example, in our trial exercise we asked the subjects to select the CRT character size that would be best for viewing under conditions of vibration. The earliest subjects looked at the relevant figure and asked us what error rate was acceptable. We selected 5% as an arbitrary but plausible value. However, no one knew what an error rate of 5% meant and several subjects pointed this out. In reading text, as compared to reading nonredundant symbols, a 5% error rate would have different impacts in each case. There was no guidance for extrapolating from the

Engineering Data Compendium entry to the users' needs. More seriously, there was potential for many users not knowing how to apply this Compendium data. We have no solutions for this within the framework of an Engineering Data Compendium. It is an issue more appropriate for the Designers' Associate. It is raised here because it will be a barrier to the use of the Engineering Data Compendium.

We have addressed the issue of extrapolation earlier (see Table 5). The major barriers were information of too general a nature, inappropriate methodology, and conflicting data. The major dimensions of difference that caused problems were the type of subject run, the way cues were presented, the way practice was conducted, data not current (i.e., techniques or, more often the hardware used, were outdated). If a report combined a variety of subjects or methods, then users could have trouble seeing how these data could be interpreted and applied to their design needs. If a research effort focused on a specific set of parameters, then users would certainly find mismatches to their decision needs. This problem can be addressed directly by helping users to make adjustments in the reported results to compensate for paradigm differences. This issue will be discussed later as a topic for future research.

We were interested in using our interview experiences to identify ways to make it easier for designers to accept and use the Engineering Data Compendium. Table 7 presents a checklist we have developed of general factors affecting the acceptance of change (Klein, 1981). It may be valuable to use it to assess the potential acceptance of the IPID products. In developing materials to help potential users learn about the advantages of the Engineering Data Compendium, some items on this checklist may be of use.

1. Agreement about the deficiencies of the current system. This is undeniable and received a strong response in our presentations.

2. Credibility of evidence that proposed change will be valuable. This remains open. Until the Engineering Data Compendium has been developed and test applications accomplished, there will not be much data on which to base judgment.

3. Relative advantage over existing system. For design engineers, this is still not clear. To the extent that design engineers do not look

for, want, or need basic research data, the Engineering Data Compendium may not offer them a great deal.

4. Complexity. The Engineering Data Compendium offers an advantage here. The human factoring has reduced the complexity level and the amount of training needed for effective use.

5. Compatibility is a plus for the Engineering Data Compendium, but not a major point.

Table 7

FACTORS AFFECTING THE ACCEPTANCE OF CHANGE

1. Agreement about deficiencies of current system.
2. Credibility of evidence that proposed change will be valuable (remedy deficiencies or extend current system).
3. Relative advantage over existing system (in terms of meeting needs, cost benefits, other dimensions).
4. Complexity -- ease of understanding and use; amount of training needed.
5. Compatibility with current system.
6. Visibility -- communication of change effects to others.
7. Guidelines for incorporating into current system; ease of adoption.
8. Boundaries drawn between what it will be useful for and where it will be irrelevant.
9. Divisibility -- ability to try on a limited basis.
10. Change agent -- and ways for phasing the change agent out.
11. Potential for reversing changes.
12. Involvement of users throughout planning and implementation.
13. Ongoing goal clarification -- what will the change look like? How will it be implemented?
14. Number of coordinating agencies and their commitment.
15. Vulnerability due to delays.
16. Supply/logistics/maintenance requirements.

6. Visibility. This will have to be ensured as the Engineering Data Compendium moves out into the field. It will not be enough to pass it on and leave it at that. In one of the early phases of this research, we prepared a survey question to examine different sources of information about the Engineering Data Compendium. In consultation with the Contract Monitor, this question was deleted in favor of other questions with greater value.

7. Guidelines. These may be necessary for certain applications. If the use of the Engineering Data Compendium was made mandatory for certain types of design efforts, guidelines would have to be developed.

8. Boundaries. These are critical. Initially, the Engineering Data Compendium may have been presented as useful for many types of problems, but now it is important to ensure that user expectations are not unrealistic.

9. Divisibility. The feasibility of customizing the Engineering Data Compendium to user specialization needs should be studied, such that shorter volumes with a higher proportion of relevant entries might be prepared to support specific subdomains of perception and human performance research.

To the extent that the Engineering Data Compendium may not be relevant to certain research questions, users should be somehow cautioned about its limitations.

One HF expert worried about design engineers who take MIL STDs and research findings at "literally" face value. He was concerned that the Engineering Data Compendium could be misused. Designers might believe they had the answers and would miss the caveats and subtle shifts in methodologies. By way of analogy, he asked if we would consider ourselves ready to design bridges if he gave us a similar compendium of phenomena related to bridge dynamics. Obviously, a compendium of useful knowledge is not sufficient for making trustworthy design decisions.

(b) Implications for Designers' Associate.

One of our goals for this research was to learn how designers make decisions, especially those involving human performance implications, so that the development of concepts for a Designers' Associate could reflect a greater sensitivity to designers needs. The major implication that we derive from our research is that design engineers are not attempting to perform decision analyses on each question that comes up. Rather, they are searching for feasible solutions by conducting their own informal research and using directly analogous or typical cases. In a few cases where published findings contributed the most to providing a solution to a design problem, the reports were used to clarify a preselected solution set, but were also subordinated to personal judgment when there were misgivings on applicability. The IPID material seemed to have great value as quick and pointed tutorials for areas that might be new to the design engineer. The information was not sufficiently detailed to resolve most questions where applied, but it helped the designer learn about the area. Therefore, a major function that a Designers' Associate can play is that of a selective tutor.

However, this is not the role envisioned for the Designers' Associate. It is planned as a decision support system, using Artificial Intelligence to help designers pose questions, access the right data, and draw the best interpretations.

One difficulty in support systems is that users frequently become so caught up in their tasks that they are reluctant to interrupt their work to seek out new information. For example, upon receiving packages of unassembled materials many people try immediately to assemble the pieces before going through the instructions. For the Designers' Associate there would be advantages to embedding it within an overall design tool, such as CAD/CAM, so that its use became more automatic.

There are some easy decisions to make in planning the Designers' Associate. 1) It should use the same terminal as CAD/CAM. 2) It should allow rapid turnaround. 3) It should have high resolution color graphics. 4) It should be user friendly. 5) It should cue the user to more information and tutorials. 6) It should allow the user to personalize his/her files.

The harder questions concern ways of presenting just the right amount of information to any specific user, of making the system applicable without translation, and of achieving intelligent data compression.

One problem is in selecting the right amount and type of information to present to designers. There is a narrow window here. For any given designer, presenting information already known is a drain on attention and counterproductive. Presenting information that is too specialized to be understood adds to confusion and further drains attention. Only the delta between what is already known and what can be readily understood is worth presenting, and this will vary for different users. Thus, the Designers' Associate has a delicate task.

An example from one of our interviews is relevant here. The designer wanted to ask a question about auditory cues. "My ambient noise level is 90 db, my voice command is 60 db, what do I need for the two together? What pitch and frequency?" Can the Designers' Associate do this? Will it be able to ask him how much variation there is in ambient noise level (and will he know the answer), how much accuracy is needed (will he be able to estimate this), what is the redundancy level of the information, etc.? In the actual case, the designer selected a woman's voice because of its novelty to the pilots. Would the Designers' Associate have helped him think of this?

The Designers' Associate can provide valuable support to engineers working out the methods for their informal studies. The engineering disciplines for the most part do not emphasize attention to the detail of the scientific method, yet experimentation is widespread in the field. The success from applying formal research techniques could no doubt be improved if guidelines were prepared which point out the limits of informal research methods and outline strategies for generalizing from ad hoc research.

(c) Future Research.

One problem worth investigating further concerns the designers' ability to extrapolate from existing data. This may be one of the important barriers to the application of basic research on human performance to questions of system design. The Designers' Associate is being prepared for design engineers, to be used for existing projects. Therefore, it is essential that these users feel comfortable in applying basic research data to operational problems.

There are theoretical arguments (e.g., Manicas and Secord, 1983) that there is a fundamental gap between scientific research and the applied needs of decision makers which constitutes a barrier to successful extrapolation. Our present research project has provided strong support for this argument. Again and again, design engineers explained to us how they cannot use basic research findings because of the problem of extrapolation. This is the fundamental problem we have found with their use of the research literature. Designers have little quarrel with the statistical reliability of the data, however, our research found instances where designers had problems accessing relevant research, as well as instances where they did not even try. One reason was that they were dubious that they would be able to use the findings. This is because the information was too specific to the experimental conditions, and because these conditions were too narrowly drawn to represent the realistic conditions where many factors interrelate.

If we were to speculate about the factors affecting extrapolation, at least three could be identified: (a) The ease of extrapolation from a given study; (b) the applied problem domain; and (c) the skill of the designer at recognizing when extrapolation is feasible and knowing how to perform it. (a) and (b) interact, since extrapolation is a process, of flowing results in one setting to another, which is affected by both settings, the original study and the applied problem. And both (a) and (b) interact with (c), the extrapolation skills of the designer, which will be more evident in familiar and well-understood domains.

The Designers' Associate is intended to support design engineers and be more than an automated data base. One prime focus of future work might be the development of a strategy for extrapolation.

What is needed is a strategy for extrapolating from a data base of basic and applied research findings along with examples of previous systems (both research and operational). One of the ways that design engineers handle human performance problems is to identify analogue systems and subsystems, use them as baselines, and use successive approximations from these to reflect the differences between analogue or comparison cases and their target case.

A research program here might first examine extrapolation within the context of the philosophy of science, through literature reviews and contacts with leading theoreticians. A second activity would be to contrast successful and unsuccessful attempts to extrapolate from Technical Reports and other types of research literature; the intent would be to identify the key dimensions that allowed or prevented extrapolation. A third activity would be to compile the leading strategies for extrapolation and evaluate these. The evaluation should be empirical as well as theoretical. The strategies should be developed for application to a set of research reports, and potential users could examine the relative advantages and disadvantages of each strategy. The results could then be used to synthesize an improved strategy. These findings could support the development of a computer-based extrapolation support system to be included with the Designers' Associate.

The Comparison-Based Prediction (CBP) approach (Klein, John, Perez & Mirabella, 1986) may be a leading candidate for an extrapolation strategy. It is designed specifically to extrapolate from comparison cases, by adjusting dimensions that are high drivers (i.e., significant differences between target and comparison cases) and it includes different strategies for using multiple comparison cases, multiple judgments, and selection of optimal comparison cases.

Design engineers who use previous systems are applying this methodology in an unstructured and informal way. In addition, time pressure, incomplete information, and other factors tend to make baselining an unsystematic process. One of the ways that a Designers' Associate can support users is to provide formats for baselining. This would provide additional benefits such as creation of audit trails documenting the basis for the estimates. In essence, we are suggesting that an analogical reasoning approach such as CBP can be useful in helping designers *extrapolate both from research reports and from* previous experiences.

There are important and straightforward tasks for applying the Designers' Associate to the problem of baselining. There is the general question of how to construct the data base and how to provide access to it. We are suggesting that the needs of the designers for an analogue-centered data base might be given equal consideration with the design tendencies of software engineers for a more typical hierarchical structure. We would suggest a data base constructed around a key feature match to similar cases. The organization scheme proposed from this perspective would be based on supporting recognition--the recognition of typicality--rather than on an analytical structure using abstract categories. This may involve a completely different way of organizing data bases.

A second topic for future research is to examine the power, as compared to the significance, of research results. We have briefly examined the feasibility of using various power tests such as eta squared, epsilon squared, and omega squared. The advantage of such tests is that they address the strength of a relationship, not the probability of obtaining it by chance. For personnel who seek applications, relationship strength is much more relevant, although rarely presented in the literature. Can these estimates be included in the Engineering Data Compendium, the Designers' Associate, or as part of a general procedure for reporting research results?

Our current assessment is that there is no one standard for reporting strength of findings. The interest in strength of relationship measures in the behavioral sciences appears cyclic. The complexity of the issue has left the field uncertain. However, we believe that the idea still has merit. We would propose it as a requirement for reporting findings in future technical reports. It will be almost impossible to use with previous studies because all of the necessary data for calculating these strength-of-effect measures are rarely included. It may be useful to develop a position paper on the current status of strength-of-effect measures and their value for future research reports.

A third topic is an assessment of AI. We have suggested the value of conducting a specific assessment of AI technology, particularly relating to expert systems, to determine how to include this technology within the Designers' Associate. This has been and is being done, but by AI advocates rather than by skeptics. We propose a more skeptical review to prepare for a worst case contingency.

To summarize, we have focused on describing the ways that designers use their experience to solve problems and have examined their use of human performance data. We feel that these were important findings for aiding the development of a Designers' Associate, and for supporting the use of the Engineering Data Compendium.

REFERENCES

Allen, T. J. (1977). Managing the Flow of Technology. MA: The Massachusetts Institute of Technology.

Frohman, A. (1968). Polaroid library usage study. Cambridge, MA: MIT Sloan School of Management, Term Paper. In Allen, T. J. (1977) Managing the Flow of Technology. MA: The Massachusetts Institute of Technology.

Janis, I. L., and Mann, L. (1977). Decision Making. New York: Free Press.

Klein, G. (1981). Checklist of factors effecting the acceptance of technological changes. Klein Associates Working Paper.

Klein, G., Calderwood, R., and Clinton-Cirocco, A. (1985). Rapid decision making on the fire ground. Technical Report No.TR-85-46-12, prepared for U. S. Army Research Institute, November.

Klein, G., John, P., Perez, R., and Mirabella, A. (1986). Comparison-based prediction of cost and effectiveness of training devices: a guide. Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Manicas, P. T., and Secord, P. F. (1983). Implications for psychology of the new philosophy of science. American Psychologist, 38, 399-413.

Rouse, W. B. (1983). On better mousetrap and basic research: Getting the applied world to the laboratory door. IEEE Systems, Man, and Cybernetics Conference.

Tversky, A. (1977). Features of similarity. Psychological Review, 84, 327-352.

APPENDIX A

Questions of Human Factors Literature

This is the total set of 132 training device design questions that arose during all of the individual and group interviews. It is a sample of the Human Factors questions that designers face.

Visual Perception

A. General Topics.

1. What visual cues are used to pilot a rotary wing aircraft in daylight? At night?
2. Will CFF be a problem in aircraft? Can the CFF be predicted? Can flicker be avoided?
3. What should the lighting requirements be in an I/O station?
4. What is the effect of off-LOS visual processing from a physically fixed position on operator comfort/performance degradation in aircrew stations?
5. What cues do Landing Signal Officers use in guiding landing aircraft on aircraft carriers? What members of the cue set need be maintained in a TD for LSOs?
6. On a TD, which controls/displays should be static and which should be dynamic?
7. Do knobs/dials have greater control/display sensitivity than button controls? What are the precision and accuracy characteristics of both?
8. What is the optimum number of control buttons to number of control adjustments relationship, e.g., less buttons, more adjustments?
9. Should a control keypad be backlit, labeled, have lighted keys for its use in a darkened TD condition? What levels of luminance will allow the maintenance of dark adaptation conditions in the user but still allow accuracy in control pad use?
10. How large should displays on a TD be so that they can be seen from the back of a classroom but not require a TD of mammoth proportions?

11. How do you light a switch's on mode when the circuit's activation shuts something off?
12. What method of presentation should be used to orient TD users to the visual angles of the real field equipment when the TD itself is in a set preselected angle?
13. How do you present complex oscilloscope waveforms in a training device, with appropriate fidelity?
14. What are the system lag time thresholds, i.e. from control input to updated system state indications, for optimum operator performance?
15. What advantages do shape coding or other tactile control configurations offer as input or display modes, versus standard visual display/controls?
16. How much resistance should be put into a dual-actionmode switch on the side of the control stick of a fly-by-wire aircraft when one mode has a critical but infrequent function?
17. What is a recommended minimum separable distance for control buttons in an aircraft?
18. How do you configure a visual display so that binoculars can be used in the training of forward observers?
19. What is the minimum lag time of control to display before negative transfer or overcontrol in TDs for aircraft piloting?
20. What type of input device/mode should be used in an I/OS for B1-B aircrew TD?
21. How should the I/OS for the above be laid out?
22. Should HUD or head-down displays be used for the above?
23. How loud do you make auditory warnings in jet aircraft cockpits?
24. What format should be used to present auditory messages in aircraft cockpits?
25. What is the gain curve for a button control, mounted sideways on the 'stick' of a fly-by-wire aircraft?

26. What are the perceptual parameters of convergence? [What information is available due to binocular vision versus monocular for depth perception? Target detection/recognition?]
27. What are the parameters of moire effects? [What symbol fill-patterns on CRT symbology will generate moire illusions? In background/foreground contexts? As a function of tasks?]
28. What is resolution capability of the retina across FOV? [How much detail is necessary to perform certain tasks and in which portions of the visual field? Does task nature effect the need for detail?]
29. What is eye movement capability in low versus high speed environments? What is the duration of the movements? Number of fixations possible? [How much information is available to the eye in the above environments will define how much information will be displayed.]
30. How much negative diopter is acceptable? [For calibration of visual displays, what are the fatigue/error thresholds of varying levels of negative diopter?]
31. How do brightness level increases compensate for loss of resolution ability - for what tasks?
32. What are the tolerance thresholds of left visual field/right visual field misalignment [for binocular displays]?
33. When does brightness level and angle of diffraction from HMD/HUD to cockpit canopy combine to initiate the Pulfrich phenomenon?

B. <u>Saccadic Suppression</u>

34. How do you predict fixation point after a saccade? Can an eye-slaved visual display system precede the eye in a saccade to arrive at the fixation point?
35. Does the eye overshoot the fixation point at the end of a saccade? Does fatigue cause this/affect this?
36. How long does saccadic suppression last? Can an eye-slaved

visual display system take advantage of this suppression to precede the eye to the new fixation point?

37. How can the saccadic suppression phenomenon be addressed in an eye-slaved TD display?
38. Once the eye has come to rest following a saccade, what visual cues are keyed on to maintain continuity of visual scene/information or perception from the previous focal point? Do normal head movements affect which cues are used and how? Does FOV affect which cues are used and how? How large a FOV is necessary to simulate the normal perceptual process in aircraft?
39. What visual information is available during a saccade?

C. <u>Depth Perception</u>

40. What resolution and FOV is needed for binocular viewing so that depth perception can be sustained?
41. What kind of information, in what detail, do pilots use to judge altitude in low altitude? At different altitudes and speeds?
42. What is the difference in focal distance in monocular viewing versus binocular viewing? How does this difference impact on monocular cueing in TDs?
43. How does resolution affect depth perception? [How much detail is necessary in a display for distance perception?]
44. How does binocular differ from monocular distance perception? [Are monocular cues sufficient for adequate depth perception?]
45. How does FOV affect depth perception? [How large does a display have to be?]
46. What are the effects of motion on depth perception? [What information is available to depth perception process when the FOV is moving - how much information needs to be displayed to preserve a sense of depth?]
47. What are the resolution, contrast, and brightness parameters of altitude perception? [How much resolution,

brightness, and contrast does a display need to preserve a sense of altitude?]

D. Angular Subtense

48. How do you build a display so that trainees perceive rate of closure in an aerial refueling task? What cues?

49. What is the relationship between real and apparent overtake rates in target acquisition?

50. How does object size in the visual field and visual field rotation affect image stability? Can size of the displayed object be manipulated to maintain consistency of perceived image size?

51. What cues are used in rate of closure perception?
[a. In aircraft environment, what cues are used to judge approach in the air-to-air refueling task?
b. What are the changes in angular subtense used by the landing signal officer when directing aircraft landings on a carrier to judge glideslope and change in glideslope - how do you represent this on a CRT screen?]

E. Color Perception

52. How is color perception affected by luminance contrast? [How bright do colors have to be maintained to preserve hue identity?]

53. What contrast is necessary to preserve cockpit display hue identity in full sunlight?

54. How is color perception affected by angular subtense? [What brightness or saturation levels need to be altered to preserve a sense of hue identity when the object in the visual field is of a certain size or is changing size?]

F. Peripheral Vision

55. How far in the periphery do pilots use information from the visual field for an aerial refueling task?

56. How much information is available from the periphery and how much is used during aerial refueling of helicopters - how large a FOV is necessary in a simulator?

57. What information is available and how is it used from the peripheral FOV when judging depth? [How much detail/what size FOV is necessary in the depiction of distance in simulations?]

58. When can information in the periphery counterbalance the absence of information in the focal area? How much information can be traded in this fashion, for what tasks? What are the minimum FOV requirements for this tradeoff to occur?

G. <u>Visual Scene Simulations</u>

59. What visual cues need be represented in digital image generation systems? For instructor pilots teaching students?

60. What pixel density in what geometric configurations should objects be depicted so that target detection can be facilitated, modeled on the real world process?

61. What should the luminance level be for an aerial refueling TD? What resolution? Should color be used? What should the FOV be?

62. For a guided weapon that aids its delivery by sending back a small rapidly changing FOV video of the environment, what resolution should the TD possess?

63. What image-generation device should be used for the above?

64. Should monochrome or textured patterns be used to paint gimbal models used to simulate in-flight formations of aircraft?

65. In CIG, what is the effect of different pixel densities on target recognition/detection? [What is the minimum/maximum number of pixels needed to portray targets so people can detect and identify them?]

66. What are the relationships of distance to viewing screen and detail density? [How close/far away can people sit to maintain a particular detail recognition level?]

67. What is the CFF for laser imagery? [Can the different CRT specifications as they are applied to laser imagery be relaxed?]

68. What are the difference and fatigue thresholds of CRT images as a function of chromaticity, luminance, FOV scaling, linearity, and geometric scaling?

<u>Workload</u>

69. How do you divide instructions between CRTs and the TD?

70. How would you measure the workload of an operator whose primary task is monitoring a CRT for infrequent signals?

71. How do you accurately determine the number of personnel needed to move a piece of field equipment, i.e. how accurate are the extrapolations of existing HF safety recommendations for the lifting/moving of equipment of given size/shape characteristics?

72. What are the optimums for displaying text information on CRTs, e.g. page and line size mixed alphanumeric information?

73. When should graphics versus pictorial displays be used?

74. What character symbol height should be used on a CRT when the background is 'busy'? Empty?

75. How do you test/measure workload of one versus two people on a helicopter mission?

76. Should there be leading zeros in keying software function calls?

77. How do you identify the workload of an ATF pilot and weapons officer?

78. What is an appropriate control and display layout and control activation sequence for a generic naval weapons delivery station TD?

79. How much information must be displayed in an I/OS so that I/O duties are proficiently performed?

80. What is the workload to be modeled in a TD of a B1-B aircrew?

81. What is the g-force blackout thresholds of experienced pilots in high speed flight?

82. How much visual detail is necessary to perform complex tasks in moving environments? [How can we prevent information overload of pilots during normal flight manuevers/weapons delivery missions?]

83. How do you estimate amount of information necessary to proficiently perform I/O duties?

84. How do different search strategies affect target recognition/detection?

Controls and Displays

85. Where is the best location in the visual field of controls/displays for quick scanning apprehension?

86. What are the guidelines of safety for labeling controls and displays?

87. When does vibration impair performance involving different size displays? Involving different control to display ratios?

88. On HMDs, what are the effects of vibration on symbol legibility?

89. On HMDs, when head movements are made, a time lag occurs before the HMD comes to rest in the new head position. When does this lag become annoying? When does performance suffer?

90. Do different aspect ratios of different displays in proximity affect human recognition of symbols/human performance? [e.g. on HUD symbology is 4:3 but cockpit displays are 1:1.]

B. IR Vision Goggles and Night Flight

91. What are the cues and effective FOV used to guide IR flight in helicopters? How large a FOV is necessary for TD for this task?

92. How low, under what conditions, can a pilot fly in IR flight? [How much visual field need be included in a training simulator?]

93. How large a FOV is available to pilots in IR flight? [FOV requirements for a TD.]

94. How does IR flight affect glideslope parameters? [How do you model glideslope on a TD for IR flight?]

95. What landing light patterns and how much lighting is required for landing a plane in IR flight?

C. CRT Usage

96. What information display format should be used in an IOS (e.g., icons, windowing, amount and type of coding to use)?

97. How large to make graphs on a display for users with little experience? A lot of experience?

98. What data entry device should be used?

99. What colors should be used in figure/background depiction on CRTs so that resolution and color contrast are maintained?

100. How large should the character/symbols be on a TD, given certain figure/background colors?

101. What are the performance impacts of CRT legibility requirements? [What is the minimum legibility on a CRT that does not impair performance?]

102. What is the tradeoff relationship between size of CRT characters, legibility, and amount of information available to the IP?

103. What are the differences between color raster systems that place light characters against dark backgrounds (U.S.A. practice) and those systems that place dark characters against light backgrounds (European practice). Can a standard approach be adopted?

104. When do such input modes as touchscreens, joysticks, and mice have advantages over each other? What is the appropriate input mode for what tasks? What are the optimal distances away from the display for each mode, and how do these distance values affect the amount of information that can be displayed?

105. What should the system lag be in presenting motion cues for a generic helicopter pilot TD?

106. What is the minimum refresh cycle on a VDT in aircraft before image stability is judged adequate?

107. What should the refresh rate be for image stability and an acceptable amount of information display on a vertical situation display for aircraft? Generally, what is the pilot-induced oscillation threshold of vertical display refresh patterns?

108. How are accommodation and focus levels related to fatigue, when using CRTs? [How far away/at what CRT resolution level need people sit before becoming fatigued from using CRTs?]

109. What are the relationships of different input modes, e.g. touchscreen, mouse, joystick, and character sizes and distance to the screen?

110. How do different information display formats, e.g. text, graphics, and pictorials, affect readability and fatigue?

111. When two or more CRTs, each possessing different refresh rates, are placed close together, can the display differences be discerned? When does the difference become irritating/cause performance impairment? At what proximity distances?

112. When can LCDs be substituted for CRTs? For what tasks? For what environments, e.g., aircraft, TD's, etc.

113. How do different fonts affect symbology legibility? [Can rounded be substituted for square or visa versa?]

114. How does size of display affect performance? [How will performance differ on 1", 5", and 6" displays of status indicators?]

115. When does the amount of information displayed exceed the operators' abilities to utilize the information (e.g. aircraft pilots)? How does human performance change when the amount of information displayed approaches the limit of human ability to adapt?

116. How is image stability affected by different sampling rates on a stroke versus raster CRT? [Should sampling rates be specified differently for raster versus stroke image generation systems?]

117. Does CRT CFF change as a function of color? Brightness? Image? Periphery?

Motion Perception

118. Does simulator sickness affect instructors?
119. What type of motion system should be used, if any, for a heavy truck TD?
120. What level should be specified for a motion cue below human perception threshold?
121. What added value does a g-seat bring when coupled with a visual display system in simulating motion? Versus g-seat alone? Versus visual display alone? Versus motion platforms and combinations of visual/non-visual and g-seat/no g-seat?
122. What are the tactile motion cues to use in g-seat modeling?
123. What is the minimum acceptable lag time in visual cueing of angular acceleration in TD's?
124. What is the vibration threshold of pilots in aircraft about to stall? Is the threshold affected by g-forces?
125. What is the minimum g-suit pressure sufficient to cue g-force pressures?
126. What vestibular cues signal the operator that a visual display is 'unnatural'.
127. What are the cue onset lag times for motion perception for the following types of motion generation systems:
    a. motion platforms and visual display changes?
    b. G-seats and visual display changes?
    c. G-seat and motion platform and visual display changes?
        - do all motion cues share similar thresholds for lag times?
128. What is the interaction of linear and angular acceleration on motion perception? [How do you construct a motion generation system that will be faithful to linear and angular acceleration inputs?]

129. What is the motion perception of g-seat cueing and visual scene cueing? [Is this type of motion generation system adequate for motion simulators? Is it qualitatively similar to the real experience of aircraft flight? Helicopter flight?]

130. What is the cause of simulator sickness?

131. Is simulator sickness affected by FOV? [How large a visual display do you build in a motion simulator to avoid simulator sickness?]

132. What are the threshold cues used by pilots to detect g-force and change in g-force? [What cues are necessary to be included in motion generation simulators to model g-force and change in g-forces for instructing pilots?

Appendix B

DESIGN QUESTIONS

These are the categories of questions raised by the designers we studied. It is an abstraction of the 132 design questions presented in Appendix A. If training device designers needed to identify material in the Engineering Data Compendium, these are the sorts of general questions with which they might begin their search.

1. Vision

1.1 CRT - General - How to determine CRT parameters?

-- Refresh rate and lag time: How to set these, effects on performance.

-- Difference between using light characters/dark background & dark characters/light background.

-- Resolution -- What is the resolution capacity of the retina across the FOV? Focal vs ambient resolution (and its impacts on presentation of details)? Information tradeoffs for presenting material in the fovea vs. the periphery?

-- Detail -- Various aspects of eye movements and how these drive the placement of greater and lesser detail. Issues of saccadic eye movements.

-- Luminance -- How to set luminance level of CRT?

-- Brightness vs. acuity tradeoffs.

CRT - Alphanumeric text - How to portray alphanumerics on a CRT?

-- Character size -- How large? What effects to consider? E.g. background, LOS, resolution, vibration, optimal page & line size. How do these interact?

-- On HMD, effect of vibration on text legibility.

CRT - Symbols - How to portray symbols on a CRT?

-- What is the target detection threshold? How does it relate to pixel density? What are the optimal target shapes? How is target stability affected by object size and visual field rotation?

-- Detecting target depth -- What cues are needed? Parameters of convergence? Binocular vs. monocular depth cues? Effect of FOV (including peripheral cues) on depth perceptions? Effect of motion, and of resolution? Difference in focal distance for monocular versus binocular.

CRT - Portrayal of Dynamics - How to portray dynamic phenomena?

-- Closure rate: what are clues to closure?

-- Real vs. apparent overtake rates.

-- Relevant features needed for stable formation flying.

-- Cues for estimating glideslope (for judging aircraft landings on a carrier).

-- Cues to judge low altitude, such as resolution, contrast, brightness, at different altitudes and speeds.

-- Guidelines for presenting g-force blackout during high acceleration flight.

CRT - Organization - How to organize the information on a CRT?

-- Formatting guidelines.

-- Use of color cues as an organizational dimension.

-- Guidelines for assignment of text, graphics, and symbology to different display parameters, graphic versus pictorial displays.

CRT - Anomalous Effects

-- CFF effects; CFF related to laser imagery.

-- Motion sickness (for trainees and for instructors).

-- Illusory parallax.

-- Eyestrain, as a function of: chromaticity, luminance, FOV scaling, linearity, geometric scaling, accommodation and focus.

-- Moire effects: What CRT symbol fill patterns will produce Moire illusions?

-- Color hue shifts, as a function of luminance changes, as a function of object size changes.

-- For binocular displays, tolerance for misalignment of left vs. right visual fields.

-- Pilot induced oscillation, as a function of refresh rate.

1.2 Night Vision

-- How to illuminate buttons in dark without affecting dark adaptation of trainee?

-- FOV needed to train night flying?

-- How does IR flight affect glidescope parameters?

1.3 Vision and TD layout

-- FOV, including extent of visual presentation into the periphery, number of CRTs needed. Example, FOV needed for training helicopter refueling in a simulator.

-- Line of Sight - Optimal vs. marginal LOS for trainees and instructors. Effect of off-LOS viewing on comfort and performance level.

-- Proximity - How does proximity affect perception, e.g., HUD symbology is 4:3 but cockpit displays are 1:1.

-- Proximity - Relation of distance to viewing screen and detail density.

-- With binocular displays, where to place parallax-related cues such as a reticle, at the distal or proximal stimulus?

-- Illumination levels for IOS.

1.4 Distortion of Optics System -- How to measure, how to determine effects.

1.5 Visual vs. Vestibular Cues

-- Vestibular cues that signal an operator that a visual display is unnatural.

2. Motion

2.1 Presentation

-- Use of g-seats and g-suits. Advantages, value for training of different configurations of vision, motion platform, g-seat and g-suit.

-- Effects of refresh rates.

-- Tactile motion cues to use in g-seat modeling.

2.2 Detection

-- Thresholds for motion detection.

-- Threshold for g-suit pressure to cue g-force pressures.

-- Threshold for detecting vibration preceding stall (and effect of g-forces.)

2.3 Tolerance

-- Tolerance for G forces, during training.

-- Tolerance for G suit pressure to simulate G forces.

3. Audition

   -- Guidelines for loudness of auditory cues.

   -- Format for presenting auditory cues.

4. Workload

   -- General procedures for measuring workload.

   -- How to estimate amount of information that will overload the IOS operator.

   -- How to measure workload for operator whose primary task is monitoring a CRT for infrequent signals.

5. Controls

   5.1 Sensitivity

   -- Amount of resistance to load into a switch.

   -- Gain curve for button controls.

   -- Advantages of shape coding or other tactile cues.

   5.2 Spacing

   -- Separation between control buttons.

   -- Control layout, especially for IOS.

   -- Best position on control panel for quick location of key controls.

   5.3 Entry methods

   -- How to enter data into a computer - pro's and con's of different techniques - touchscreens, joysticks, mice. Effects of input mode on required distance from CRT.

   -- Relative advantages of knob by button controls.

   5.4 Safety

   -- Guidelines for safety in labelling controls

6. Anthropometry

7. Miscellany

   -- Ways of representing malfunctions.

   -- On a HMD, what is the lag threshold for HMD casing to follow head movements?

Appendix C

21 RECOMMENDATIONS FOR DESIGNERS' ASSOCIATE

These are the different suggestions made by the designers for developing an effective decision support system.

1. Answers/data explained at inquisitor's level.
2. Data should be accessible.
3. Examples first, then general principles.
4. User friendly.
5. Menu sequences of what to do for certain goals.
6. Key word querying capability.
7. AI for misspellings, synonyms.
8. Caveats for applying the designers' associate especially for non-HF expert. Examples of suitable data applications.
9. Window displays.
10. Query language should be naturalistic; commands should have clear logical meaning.
11. Sentence completion type of entry capability, e.g. "my ambient noise level is 90 db, my voice command is 60 db, what db is needed for an auditory warning? What pitch? What frequency?
12. Data base operating system should have 'novice' through 'expert' levels of interaction, display (e.g., substitution and explication of acronyms).
13. Designated function keys.
14. Avoid light pens, touch sensitive and low resolution displays.
15. Immediate screen display response to inputs.
16. Immediate screen display response to inputs.
17. Relational data system to make it easy to move data around.
18. Transform as well as access data.
19. Repository of past projects, indicating how things were done, problems encountered, similar projects.
20. Concern about maintaining resolution quality of compendium graphics. May need special CRT.
21. Use of trees for easier search patterns.

U.S. GOVERNMENT PRINTING OFFICE 1987 - 548-002/60215

END

DATE

FILMED

DEC.

1987